



Spanning Tree Protocol: Evolução



“Fast Switching”

PVST – Per VLAN Spanning Tree (Cisco)

PVST+ – Per VLAN Spanning Tree plus (Cisco)

MSTP – Multiple Spanning Tree Protocol - IEEE 802.1s

RSTP – Rapid Spanning Tree Protocol - IEEE 802.1w



**“Understanding Rapid Spanning Tree Protocol (802.1w)”, Document ID:
24062, Cisco**

Evolução do STP



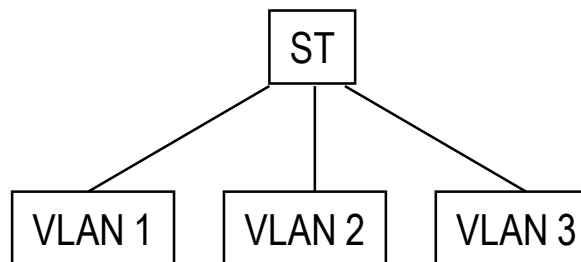
- Problema do STP:
 - Lento a convergir para a topologia activa
- Soluções:
 - *Fast Switching e Rapid Spanning Tree Protocol (RSTP)*
- Problema do STP:
 - Não funciona bem com as VLAN
 - Com topologias de VLAN diferentes, o STP ao nível do porto físico não é eficiente
- Soluções:
 - PVST, PVST+, ST Domains e MSTP

Evolução do STP

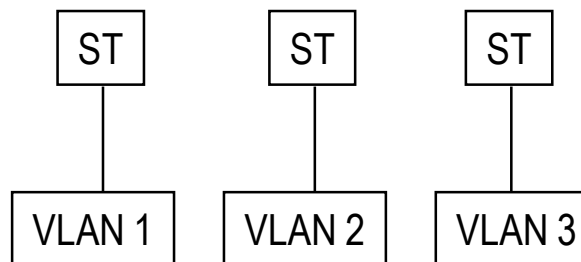


ST – *Spanning Tree*

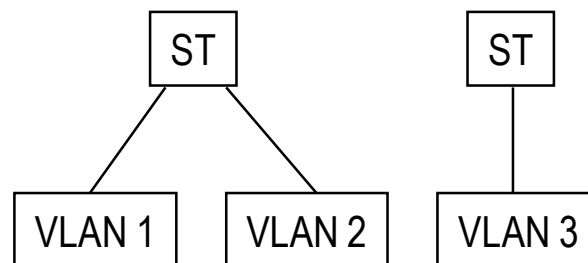
Spanning Tree Protocol
Rapid Spanning Tree Protocol



Per VLAN Spanning Tree Protocol
Per VLAN Spanning Tree Protocol+



Multiple Spanning Tree Protocol



Normas Relacionados com STP



- 802.1D - 1990, actualizado em 1998
 - *Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges*
 - *Transparent Bridging, Spanning Tree Algorithm and Protocol, Source-Routing Transparent (SRT) Bridges*
- 802.1t - 2001
 - Correções técnicas e editoriais ao 802.1D-1998
- 802.1w - 2001
 - *IEEE Standard for Information Technology -Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Common Specifications - Part 3: Media Access Control (MAC) Bridges: Amendment 2 - Rapid Reconfiguration*
 - Reconfiguração rápida do STP (RSTP)
- 802.1s - 2002
 - Múltiplos STP num *switch/bridge* com VLANs (MSTP)

Portas *Fast Switching*



- Objetivo
 - Permitir que portas que se sabe à partida que não vão provocar *loops* possam passar ao estado *forwarding* mais rapidamente
 - Ex.: Ligações a máquinas ou *routers*, *uplinks* sem redundância
- Funcionamento
 - A porta configurada como *FastSwitch* (*Fast Learn*, etc) não executa o protocolo *Spanning Tree* (STP)
 - Pode voltar a executar o STP se receber um BPDU
- Alternativa (menos segura)
 - Fazer *disable* do STP nos portos que se sabe que não vão provocar *loops*

RSTP - *Rapid Spanning Tree Protocol*



- Definido na norma IEEE 802.1w de Junho de 2001
- Objetivo:
 - Organizar os segmentos da rede em árvore num tempo na ordem das dezenas de milissegundos
- Características:
 - Suporta *Bridges/Switches* com mais de 256 portas
 - Suporta *FastSwitching / Forwarding* em portas nas extremidades da rede
 - Compatibilidade com o STP
 - As portas passam para STP se receberem algum BPDU
 - As portas podem funcionar em modos diferentes (STP, RSTP)

Rapid Spanning Tree Protocol

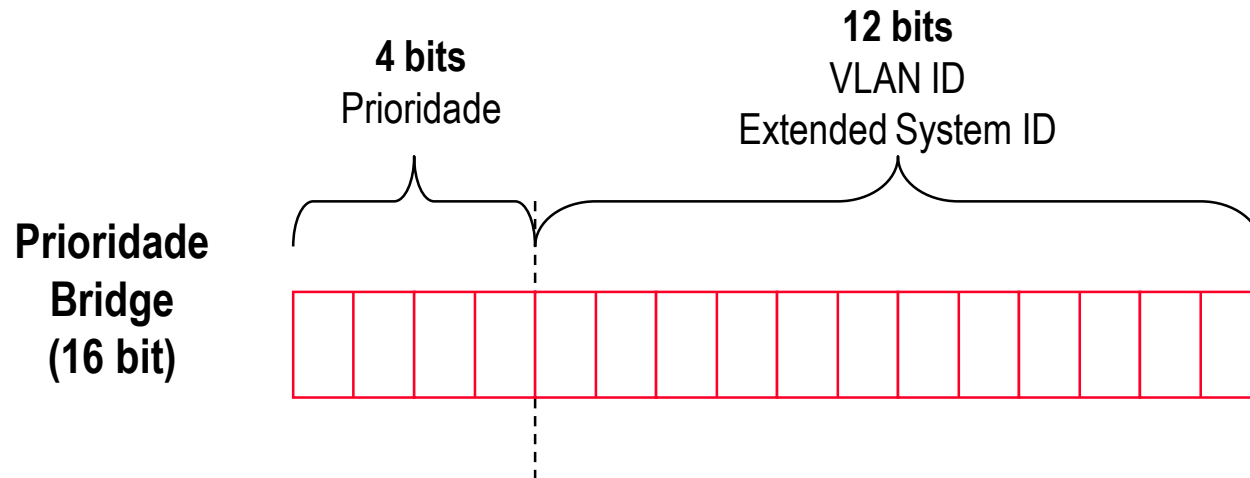
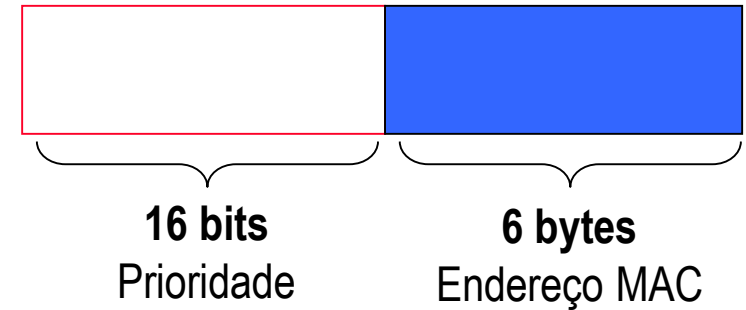


- Funcionamento:
 - Semelhante ao STP
 - Os BPDU incluem informação adicional
 - Aproveita o facto das ligações actuais serem quase sempre ponto-a-ponto

802.1s – Identificador de *Bridge* (Alterações ao STP)



- Prioridade Bridge (16 bits)
 - 4 bits para prioridade *bridge*
 - 12 bits para o **identificador da VLAN**



Valores Configuráveis:
[0..61440]
Incremento:
4096
Valor por omissão:
32768

Alterações nas *flags* das mensagens BPDU



Flags definidas no STP:

0 - TC – *Topology Change*

7 - TCA – *Topology Change Acknowledge*

Flags definidas no RSTP:

0	1	2	3	4	5	6	7
---	---	---	---	---	---	---	---

0 – *Topology Change*

1 – *Proposal*

2, 3 – *Port_role* → { 00 Unknown

01 Alternate/Backup

10 Root

11 Designated

3 – *Learning*

4 – *Forwarding*

6 – *Agreement*

7 – *Topology Change Acknowledge*

C-BPDU

Nº
Octetos

<i>Protocol Identifier</i>	2
<i>Protocol Version Identifier</i>	1
<i>BPDU Type</i>	1
<i>Flags (TCA, Reserved, TC)</i>	1
<i>Root Identifier</i>	8
<i>Root Path Cost</i>	4
<i>Bridge Identifier</i>	8
<i>Port Identifier</i>	2
<i>Message Age</i>	2
<i>Max Age</i>	2
<i>Hello Time</i>	2
<i>Forward Delay</i>	2

Protocol Version = 2
BPDU Type = 2

Quem gera os BPDU?

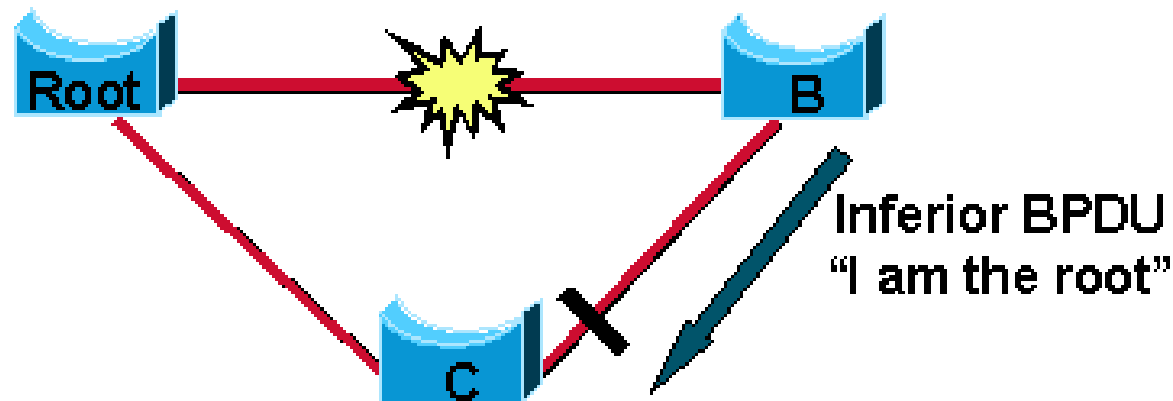


- Ao contrário do STP em que numa topologia estável apenas a *root* é que gera os BPDU, **no RSTP todos os switches geram e enviam um novo BPDU a cada <hello-time> segundos (2 s por omissão).**
- Se uma *bridge* não receber 3 BPDU seguidos assume que perdeu a ligação direta com os seus vizinhos (funcionam como *keep alive*).

BPDU inferiores



- Quando uma *bridge* recebe informação inferior (BPDU piores) da sua *bridge root* ou *designated*, ela aceita-a e substitui o BPU guardado anteriormente.

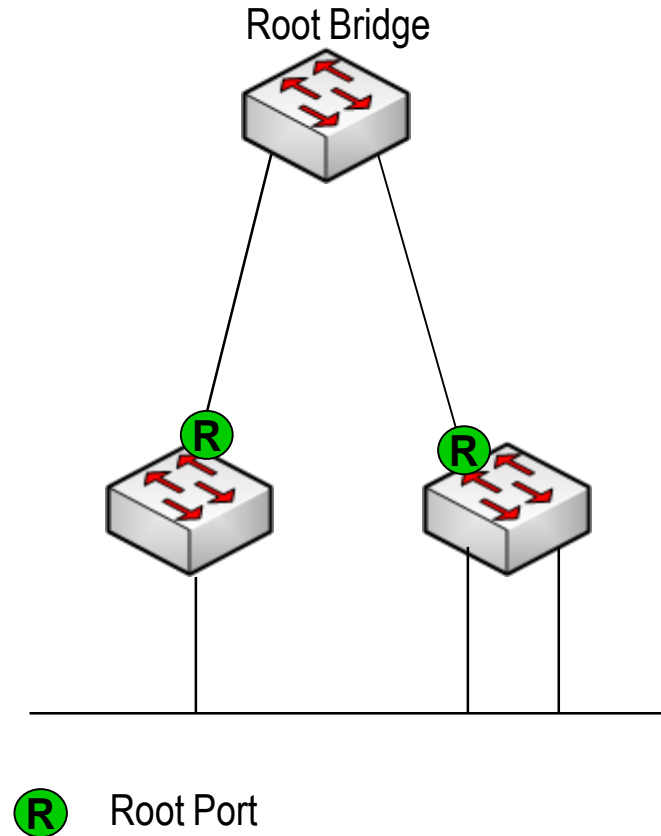


- No caso da figura, porque a *bridge* C ainda sabe que a *root* está “viva” e bem, ela envia imediatamente um BPDU à *bridge* B que contem informação sobre a *bridge root*. Como resultado a *bridge* B não envia os seus próprios BPDU a dizer que é a *bridge root* e aceita a porta que a liga à *bridge* C como a nova *root port*.

Root Ports



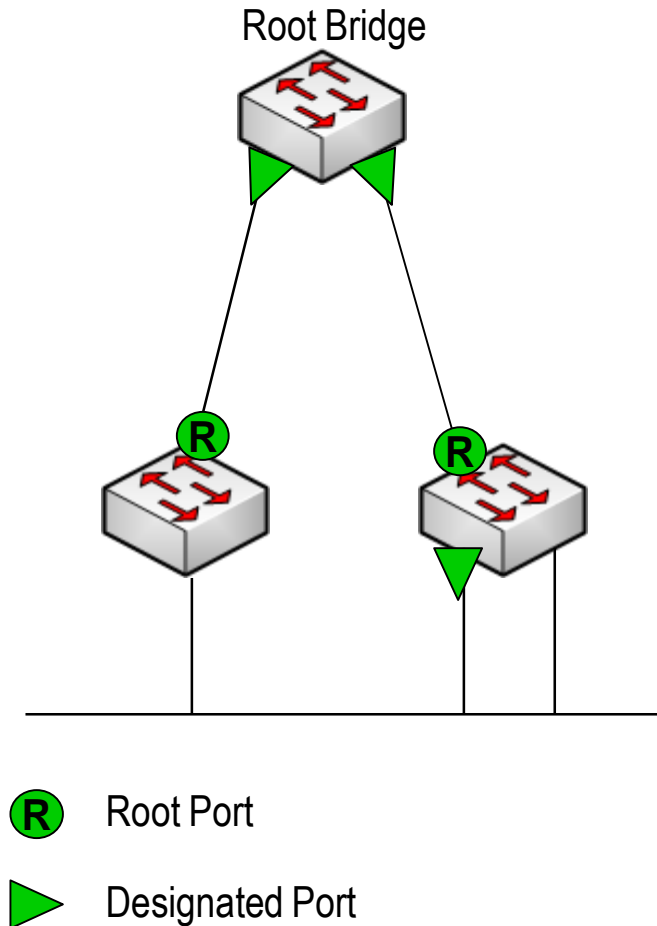
- As portas das *bridges* que não são *root bridge* e que recebem o melhor BPDU da *root bridge* são *root ports*
- A *root port* é a porta mais próxima da *root bridge* em termos de custo (*Root Path Cost* - RPC)
- O processo de escolha *root ports* é igual ao STP



Designated Ports



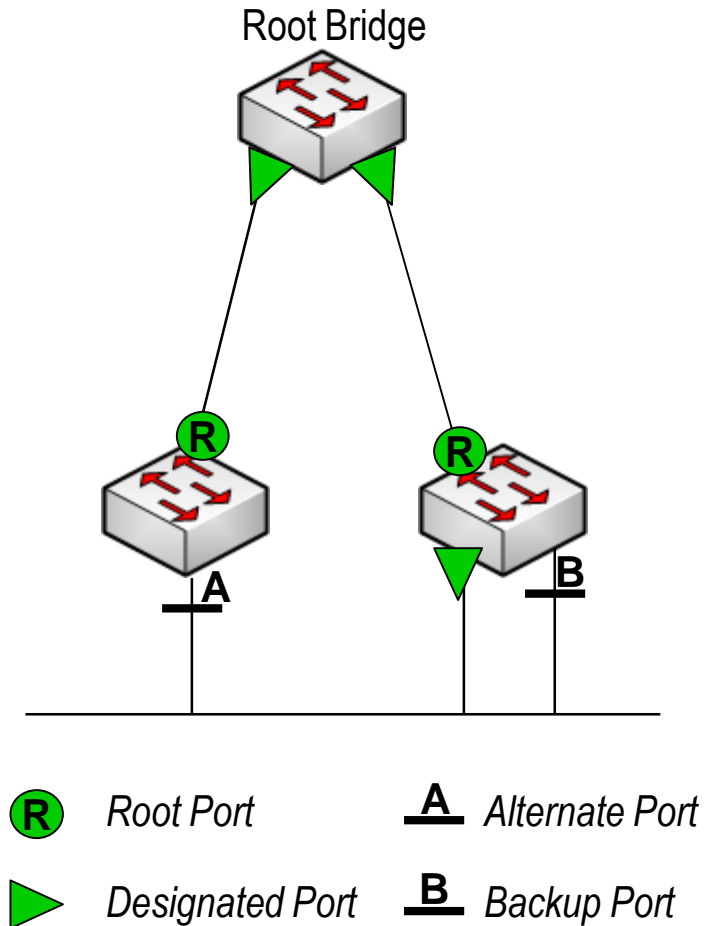
- As *Designated Ports* enviam os BPDU e as restantes tramas para o segmento a que estão ligadas
- Para cada segmento apenas deve haver um único caminho até à *root bridge*
- Todas as *bridges* ligadas a um segmento escutam os BPDU umas das outras e acordam em qual a melhor *bridge* para enviar os BPDU e outros dados para o segmento (*designated port*)
- Processo igual ao do STP



Alternate e Backup Ports



- Estes dois estados correspondem ao estado *blocked* do STP
- Uma porta bloqueada recebe os BPDU e processa-os
- As *Alternate Port* oferecem um caminho alternativo para a *root bridge*
- As *Backup Port* oferecem um caminho alternativo para o segmento (duas portas, mesma *bridge*)



Funções das Portas - Resumo



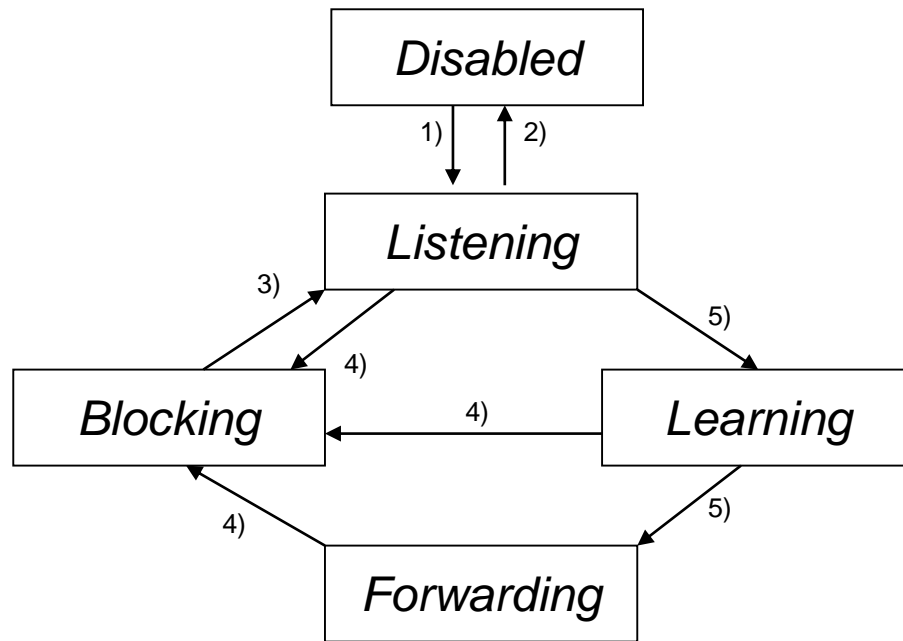
Port Role	Description
<i>Root port</i> (Igual ao STP)	A <i>root port</i> é a porta que está mais perto da <i>root bridge</i> para os <i>switches</i> que não são <i>root</i> . Apenas pode existir uma <i>root port</i> por cada <i>switch</i> . A <i>root port</i> assume o estado <i>forwarding</i> numa topologia ativa estável. Num <i>switch</i> a porta que recebe o melhor BPDU será a <i>root port</i> .
<i>Designated port</i> (Igual ao STP)	Cada segmento tem uma porta <i>designated</i> para esse segmento. Numa topologia ativa o <i>switch</i> que recebe as tramas do segmento mais próximo da <i>root bridge</i> . A <i>designated port</i> assume o estado <i>forwarding</i> . Todos os <i>switches</i> ligados a um segmento disputam entre si o estado <i>designated</i> escutando todos os BPDU e determinam entre si qual o <i>designated port</i> do segmento. <u>A porta que enviar o melhor BPDU para um segmento será a <i>designated port</i> desse segmento.</u>
<i>Alternative port</i> (Non-Designated Port / <i>blocking</i> em STP)	Uma <i>alternative port</i> é uma porta que oferece um caminho alternativo para a <i>root bridge</i> . Recebe BPDU melhores que os seus. A <i>alternative port</i> assume o estado <i>discarding</i> numa topologia ativa estável. A porta <i>alternative</i> acontece em <i>switches</i> não <i>designated</i> do segmento e transita para <i>designated</i> caso o <i>switch designated</i> falhe.
<i>Backup port</i>	Uma <i>backup port</i> é uma porta adicional num <i>switch</i> que é <i>designated</i> de um segmento e corresponde a uma ligação redundante para esse segmento. A <i>backup port</i> assume o estado <i>discarding</i> numa topologia ativa. Uma porta de <i>backup</i> recebe BPDU melhores de outra porta do mesmo <i>switch</i> . A <i>backup port</i> tem um identificador <i>port ID</i> maior que a <i>designated port</i> .

Estados das Portas em RSTP

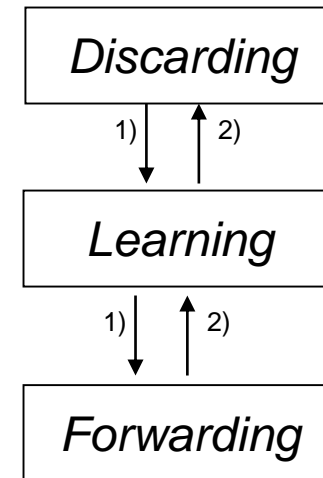


Estado da Porta	Descrição
Discarding	<ul style="list-style-type: none">• Pode acontecer quando existe uma topologia ativa ou na fase de sincronização de topologia.• Impede o envio de tramas para evitar os <i>loops</i> de nível 2.
Learning	<ul style="list-style-type: none">• Pode acontecer quando existe uma topologia activa ou na fase de sincronização de topologia.• Aceita as tramas para preencher as tabelas de encaminhamento e evitar a propagação das tramas para endereços ainda não conhecidos.
Forwarding	<ul style="list-style-type: none">• Apenas acontece quando existe uma topologia activa.• As portas Forward determinam a topologia.• Depois de uma alteração de topologia ou durante a sincronização, o encaminhamento das tramas apenas acontece depois de haver acordo de topologia.

Estados das Portas e suas Transições: STP/RSTP



STP



RSTP

Relação entre os estados das portas



Relação entre o estado operacional das portas e os estados STP e RSTP

Estado Operacional	Estado STP	Estado RSTP
Disabled	Disabled	Discarding
Enabled	Blocking	Discarding
Enabled	Listening	Discarding
Enabled	Learning	Learning
Enabled	Forwarding	Forwarding

Estados e Funções das Portas (STP vs RSTP)



• Estado - STP (802.1D)

- Disabled
- Blocking
- Listening
- Learning
- Forwarding

• Estado - RSTP (802.1W)

- Discarding
- Discarding
- Discarding
- Learning
- Forwarding

• Função - STP (802.1D)

- Root Port
- Designated Port
- Blocking
- Blocking

• Função - RSTP (802.1W)

- Root Port
- Designated Port
- Alternate Port
- Backup Port

• Estado Forwarding

- Root Port
- Designated Port

• Estado Discarding

- Alternate Port
- Backup Port

RSTP – Custos das portas (por omissão)



Port Speed	Link Type	Path Cost 802.1D-1998	Path Cost 802.1w
10 Mbps	Half Duplex	100	2,000,000
	Full Duplex	95	1,999,999
	Aggregated Link	90	1,000,000*
100 Mbps	Half Duplex	19	200,000
	Full Duplex	18	199,999
	Aggregated Link	15	100,000*
1000 Mbps	Full Duplex	4	20,000
	Aggregated Link	3	10,000*

* This path cost is correct where there are two ports in an aggregated link. However, if there are more ports in the aggregated link, the path cost will be proportionately lower. For example, if there are four ports in the aggregated link, the 802.1w path costs will be: 500,000 for 10 Mbps, 50,000 for 100 Mbps, and 5,000 for 1000 Mbps. The 802.1D-1998 path cost values are not affected by the number of ports in an aggregated link.

Transição rápida para o estado Forwarding

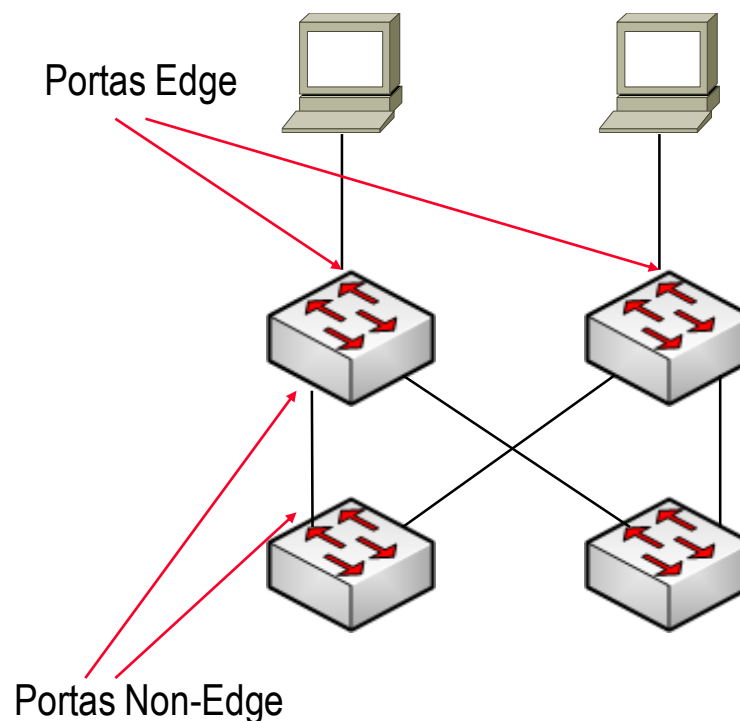


- A rápida transição é a característica mais importante introduzida pelo 802.1w.
- O novo RSTP é capaz de ativamente confirmar que uma porta pode transitar em segurança para o estado *Forwarding* sem ter de se basear em qualquer *timer*.
- Existe agora um mecanismo real de *feedback* que atua entre *bridges* RSTP. De maneira a obter uma convergência rápida de uma porta, o protocolo confia em duas novas variáveis:
 - Portas tipo *edge*
 - Portas tipo *link*.

Portas Edge e portas Non-Edge



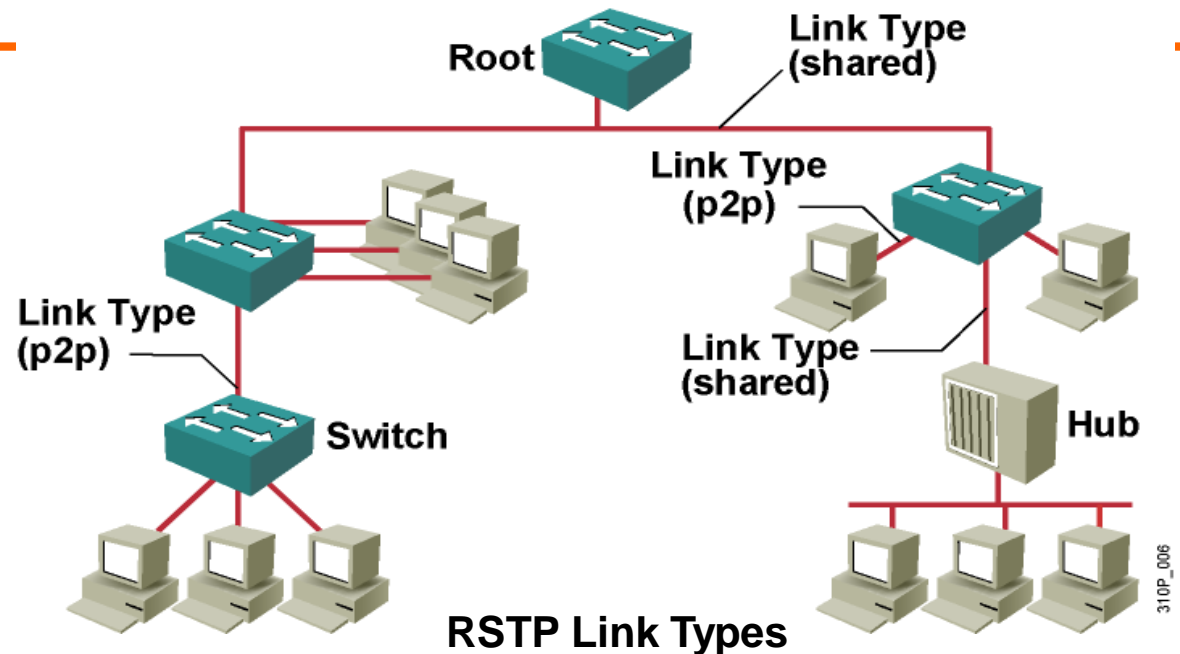
- As portas *Edge* nunca terão outros *switches* ligados a elas
- Transitam imediatamente para o estado *Forwarding*
- Têm funções similares ao *PortFast da Cisco*
- São configuradas para aceitarem o comando *spanning-tree portfast*
- Uma porta *edge* que receba um BPDU passa a porta *Non-Edge*



Tipos de *Link* RSTP

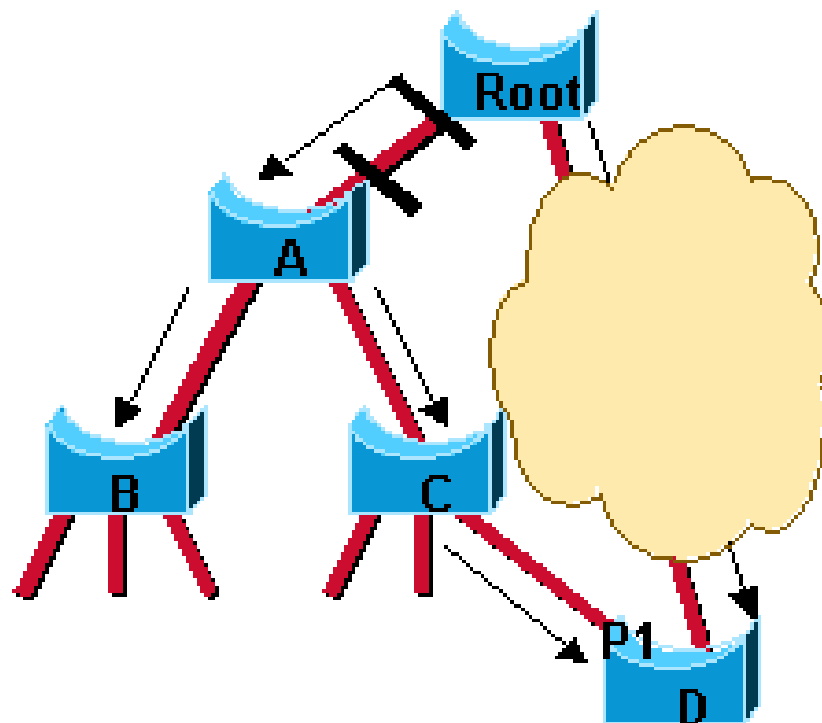


- O tipo de *link* determina o papel ativo que a porta desempenha para transitar rapidamente para o estado *forwarding* em função de alguns parâmetros.
- Os parâmetros são diferentes para as *edge ports* e para as *non-edge ports*.
- As portas *non-edge* têm dois tipos de *link* possíveis:
 - ponto a ponto, e
 - partilhados.
- O tipo de *link* é determinado automaticamente mas pode ser alterado por configuração da porta.



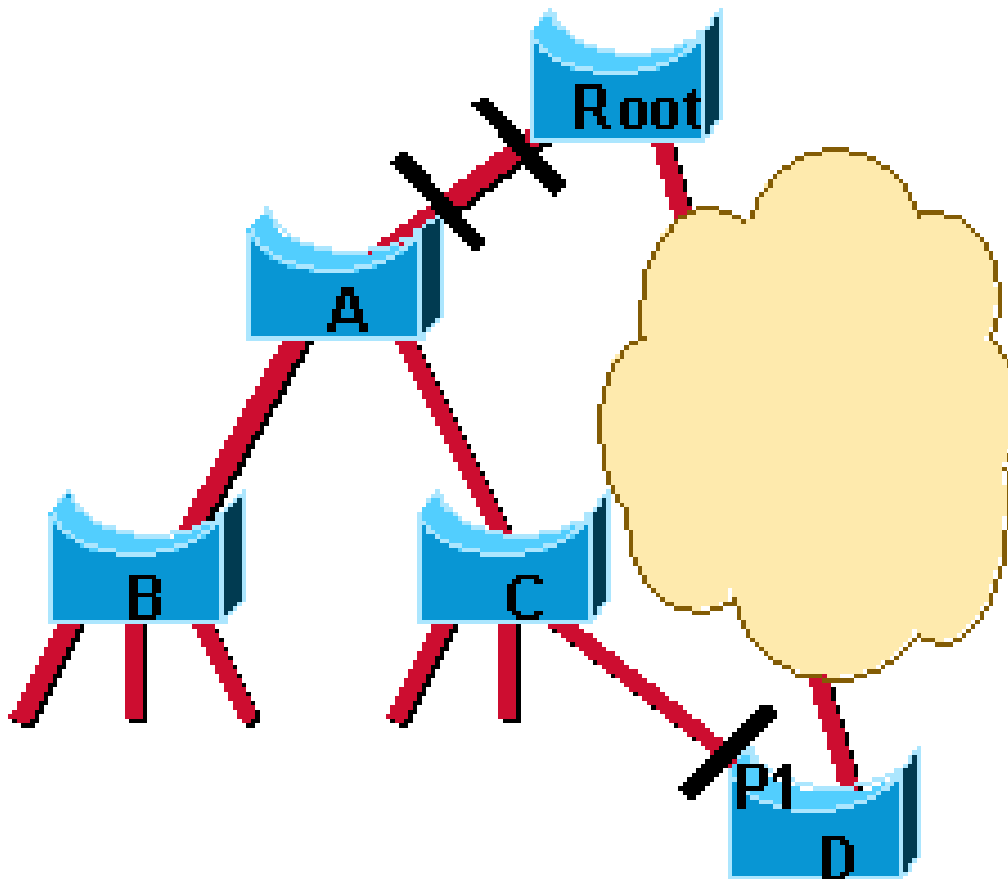
Tipo de Link	Descrição
<i>Point-to-point</i> (Ponto a Ponto)	<ul style="list-style-type: none">• A porta opera em modo <i>full-duplex</i>.• Assume que a porta está ligada a um único <i>switch</i>.
<i>Shared</i> (Partilhado)	<ul style="list-style-type: none">• A porta opera em modo <i>half-duplex</i>.• Assume que a porta está ligada a um meio partilhado onde podem existir múltiplos <i>switches</i>.

Convergência da topologia em 802.1D



Adicionado novo *link* entre a *root* e a *bridge* A. Portas que se ativem na *bridge* A ou na *root* são imediatamente colocadas no estado ***listening*** e o tráfego bloqueado. Os BPDU da *root bridge* começam a ser propagados para as folhas através de A.

Convergência da topologia em 802.1D (STP)



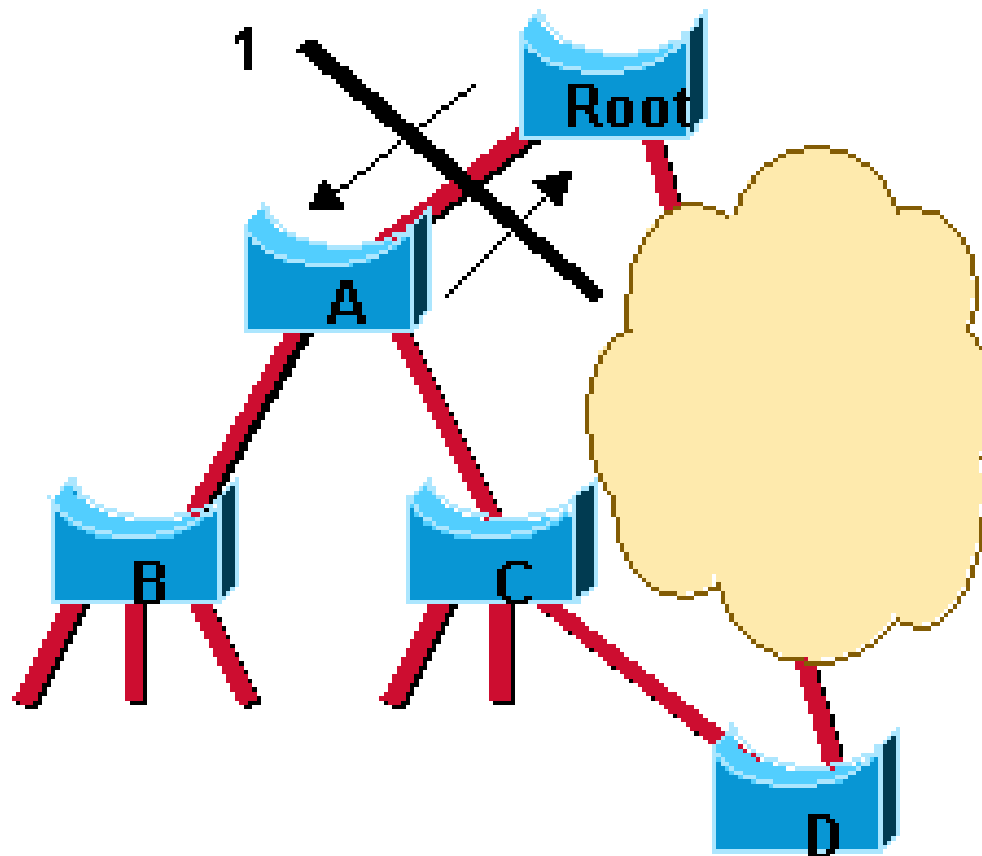
Rapidamente os BPDU da *root* chegam a D, a qual imediatamente bloqueia a sua porta P1.
A topologia convergiu mas a rede é perturbada durante $2 \times \textit{forward_delay}$ (30 segundos por omissão)

|

Convergência da topologia em 802.1w (RSTP)



Ambas as portas da ligação entre A e a *root* ficam em *designated blocking* assim que vêm acima, semelhante ao STP. Os *switches root* e A negociam. Assim que o *switch A* recebe um BPDU da *root*, bloqueia todas as portas *designated non-edge*. Esta operação designa-se Sync. Uma vez isto feito, a *bridge A* autoriza a *bridge root* a colocar a sua porta no estado de *forwarding*.

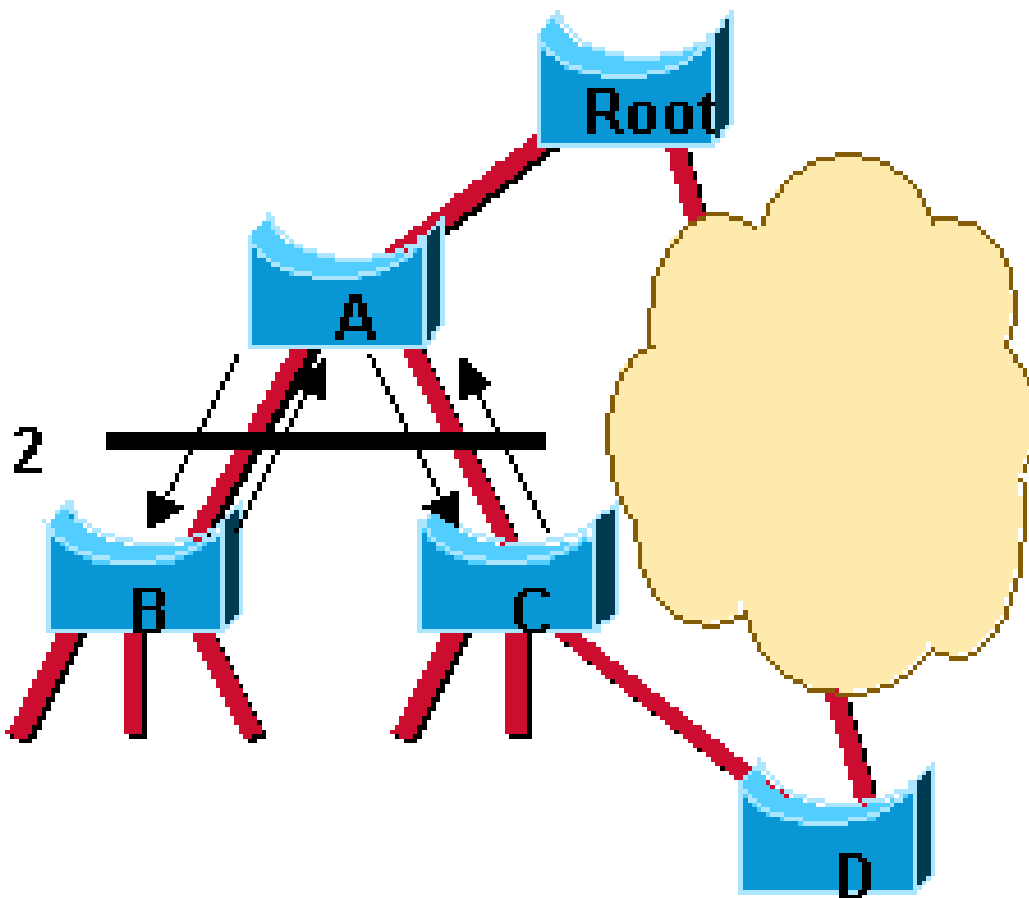


A ligação entre as *bridges root* e A é bloqueada e ambas as *bridges* trocam BPDU.

Convergência da topologia em 802.1w



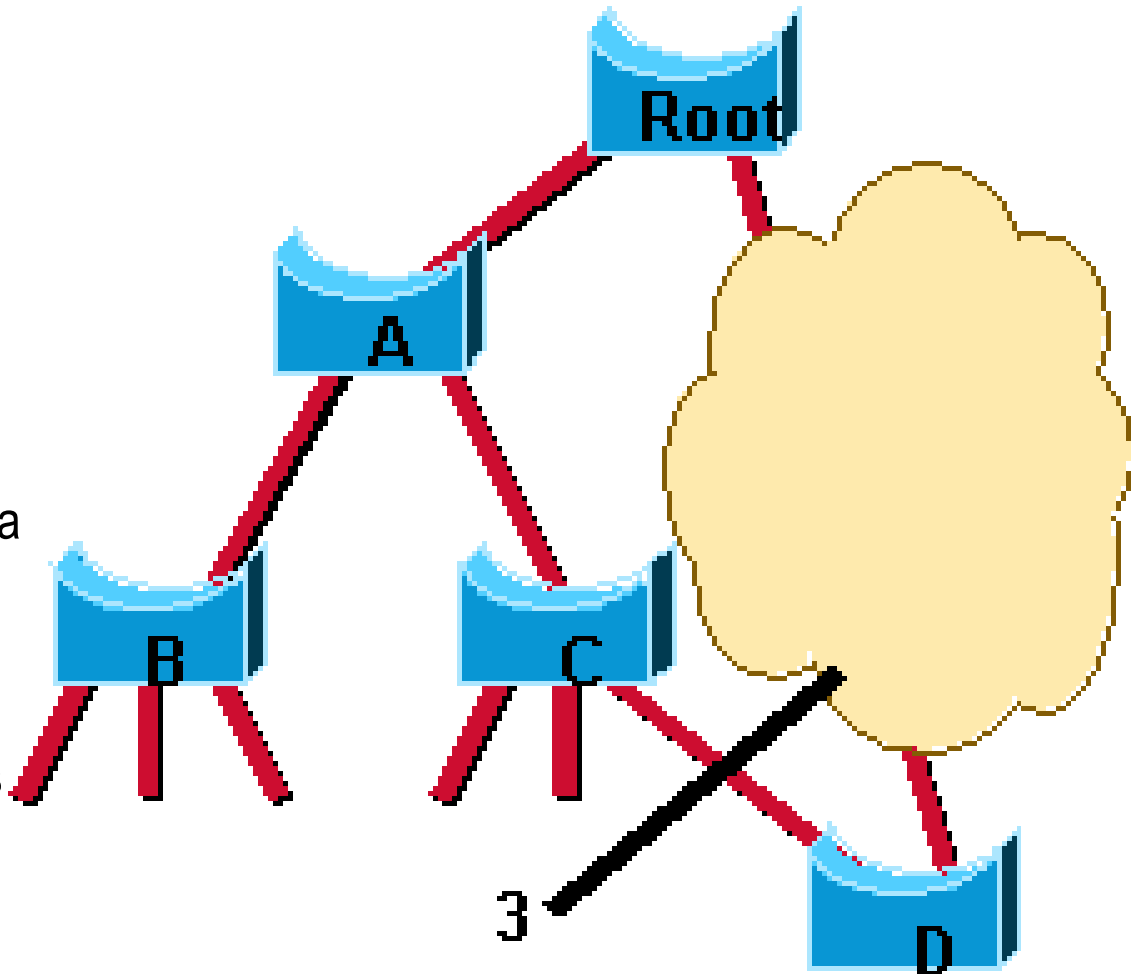
Quando o *switch* A bloquear as suas portas *designated non-edge*, a ligação entre os *switches* root e A é colocada no estado de *forwarding*.



Convergência da topologia em 802.1w (RSTP)



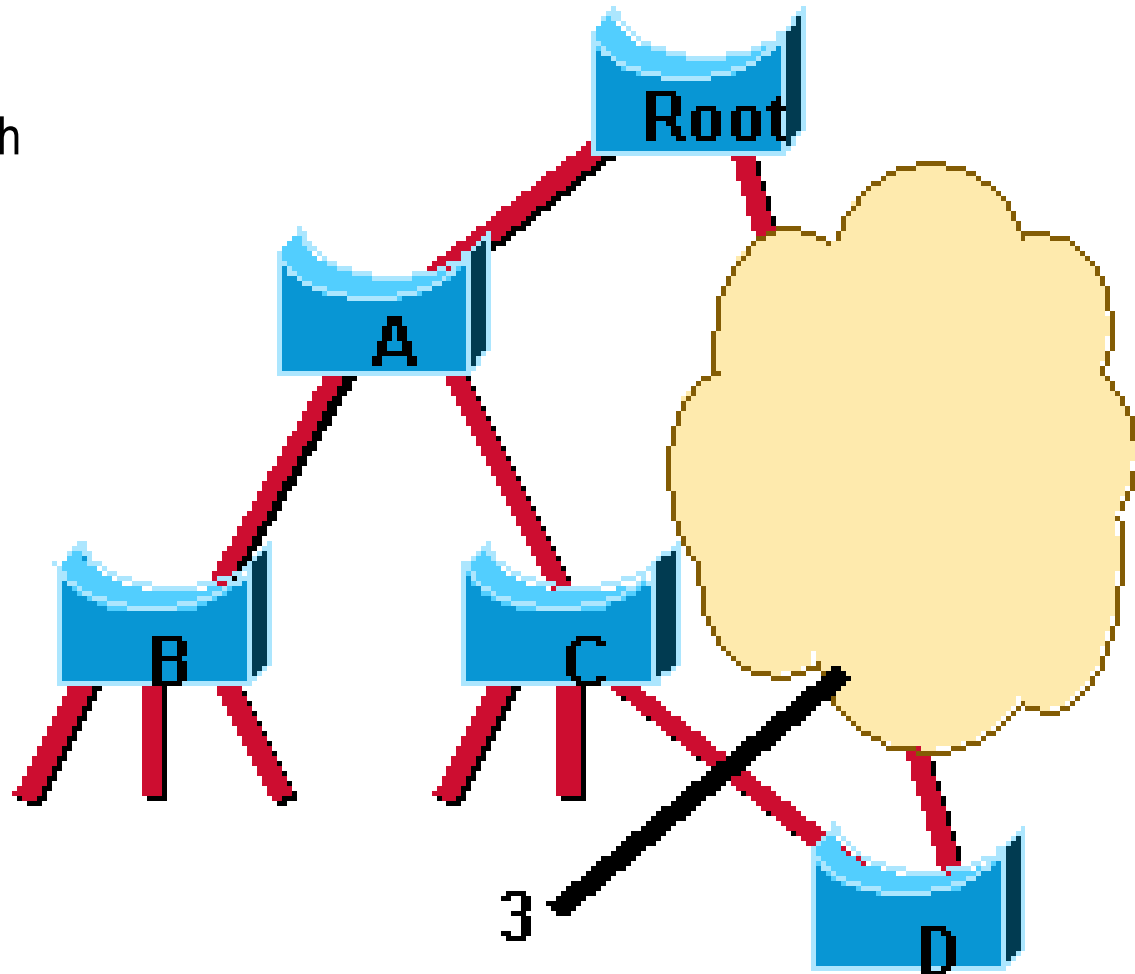
Ainda não pode existir um *loop*. Em vez de bloquear acima do *switch A*, a rede é bloqueada abaixo deste. O corte dos eventuais *loops* viaja de cima para baixo na árvore com os novos BPDU da *root* enviados via *switch A*. Nesta etapa as portas no *switch A* agora bloqueadas negociam uma rápida transição para o estado *forwarding* com as suas portas vizinhas nos *switches B e C* que também iniciam a operação Sync.



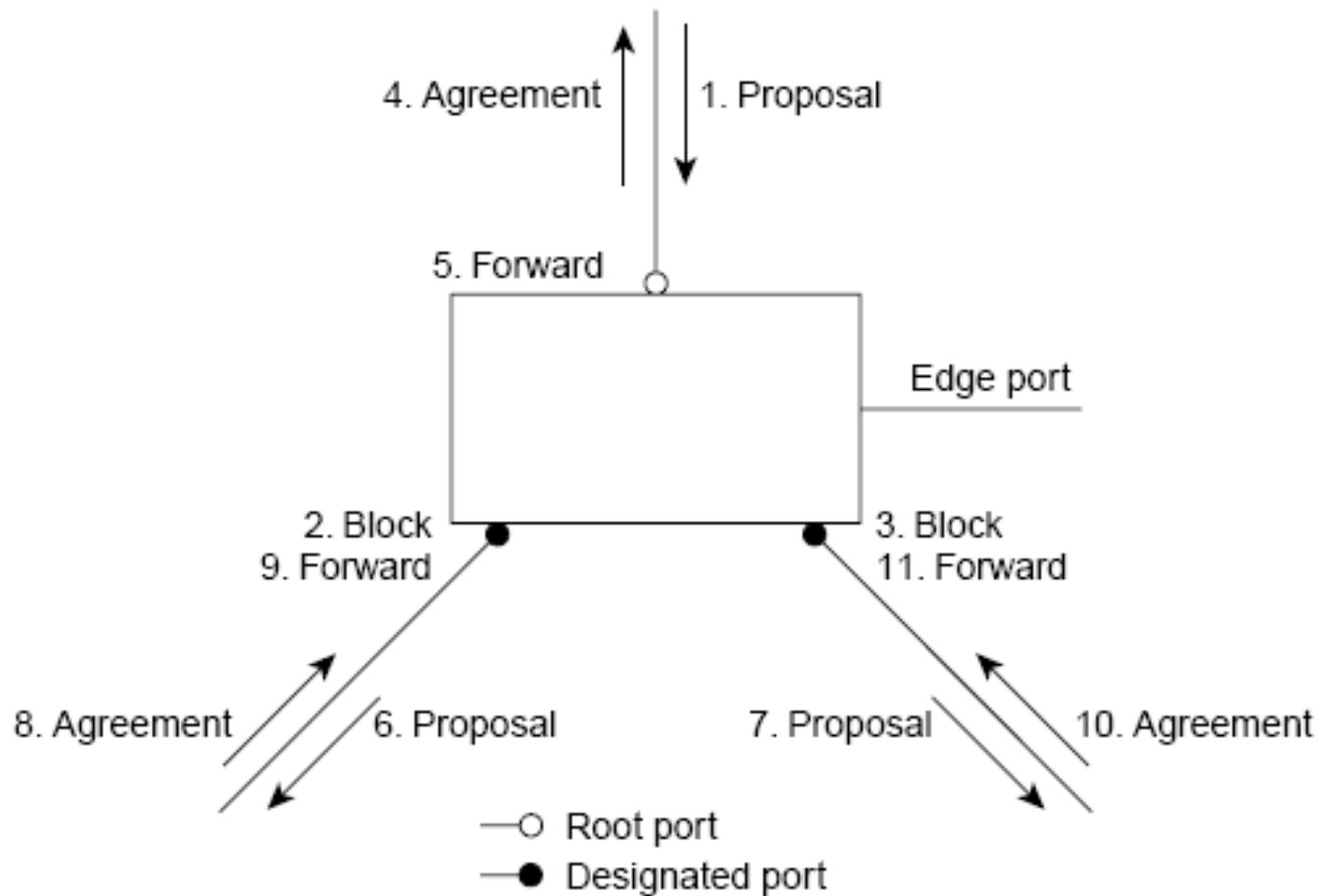
Convergência da topologia em 802.1w (RSTP)



Excluindo as portas para o switch A, o switch B apenas tem portas *edge designated*. Então, não possui portas para bloquear de maneira a autorizar o switch A a passar para o estado de *forwarding*. De maneira semelhante, o switch C apenas tem para bloquear a sua porta *designated* para o switch D.



Sequência de eventos durante convergência rápida Proposal/Agreement



Funcionamento das *flags Proposal / Agreement*



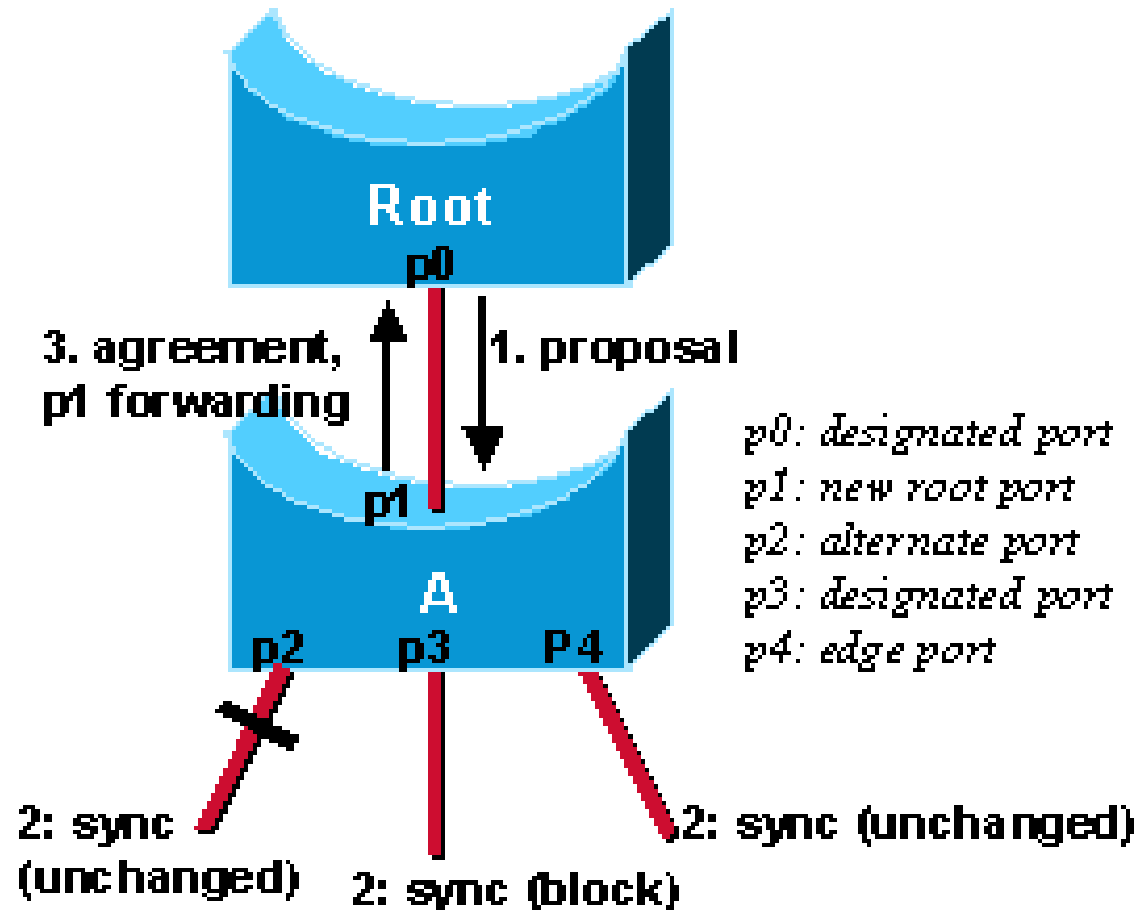
- Quando uma *Bridge* recebe uma mensagem *proposal* num dos seus portos e esse porto é seleccionado como novo *root port*, então todos os outros são sincronizados com a nova informação.
- Após confirmar que todos os portos estão sincronizados com a nova informação, envia uma mensagem *agreement* para o *designated switch* correspondente ao seu *root port*.
- As portas não *root* e não *edge* atuam de modo idêntico para decidir se ficam no estado *forward* ou *blocked* relativamente aos outros *switch*.

Alteração de Topologia: STP vs RSTP

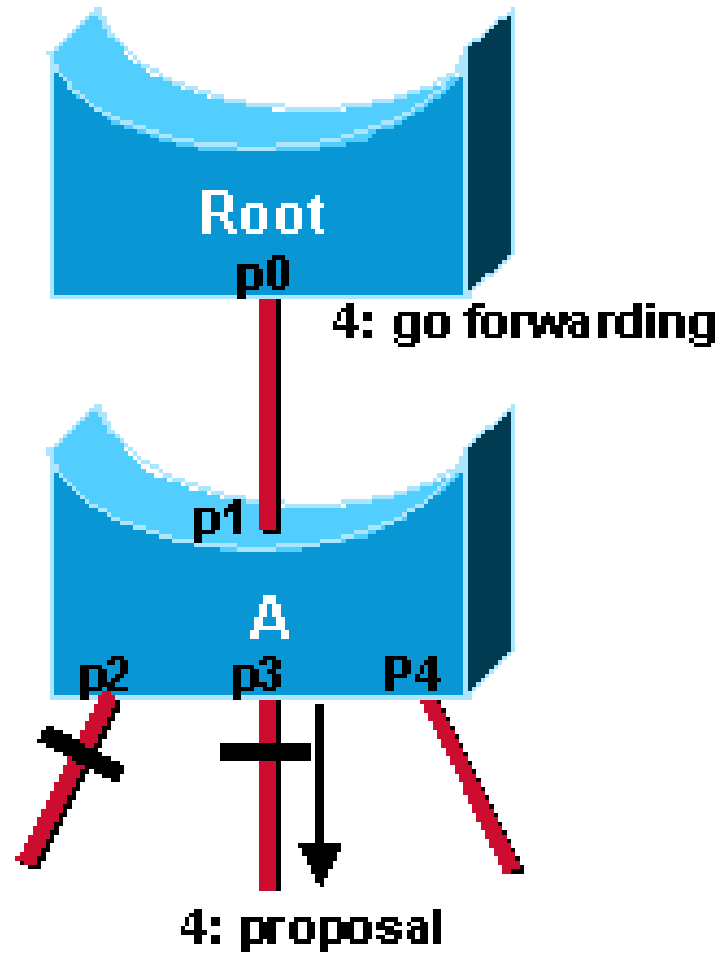


- Alteração de topologia em STP
 - Qualquer alteração de estado de uma porta gera um TCN BPDU
 - Quando uma *bridge* deteta alteração de topologia (TC) envia um TCN (*topology change notify*) BPDU para a *root bridge*
 - A *Root Bridge* envia TC BPDU (BPDU de configuração)
 - Quando um *switch* recebe estes BPDU diminui o *aging time* para o *forward delay* (para refrescar a tabela de endereços MAC)
- Alteração de topologia em RSTP
 - Apenas as portas *non-edge* que passarem a *forwarding* originam os TC BPDU
 - A falta de conectividade não se considera mudança de topologia. Neste caso um porto passar para o estado *blocking* não gera TC BPDU
 - Os *switches* limpam a tabela MAC e passam ao estado *forward* quase imediatamente
 - A propagação de mudança de topologia é um processo de um passo
 - Quem inicia a mudança de topologia envia continuamente a informação em vez de ser apenas a *Root* a fazê-lo

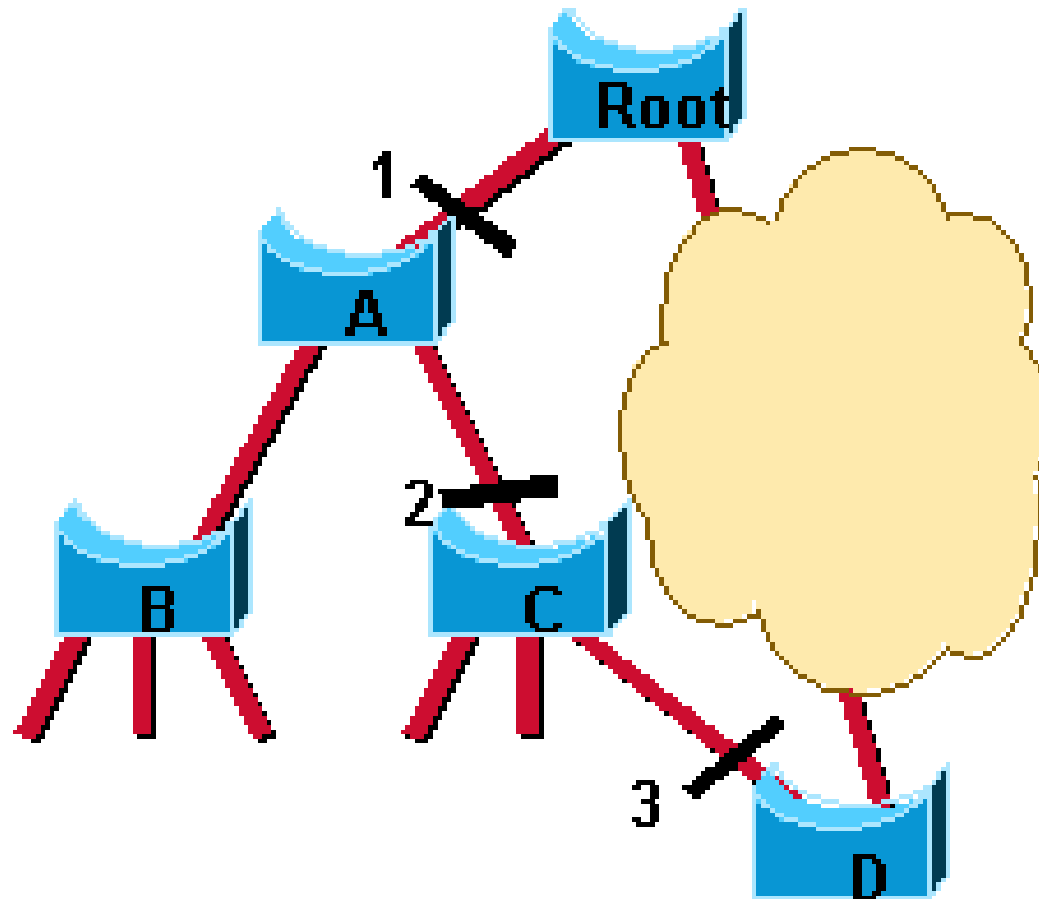
Proposal/Agreement Sequence



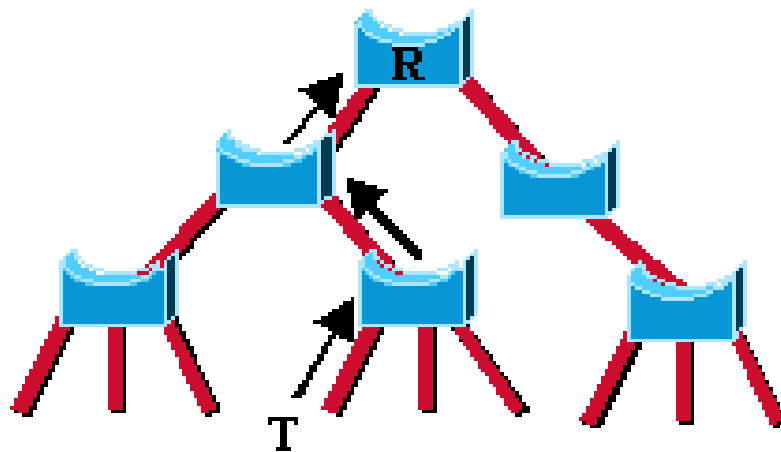
Proposal/Agreement Sequence



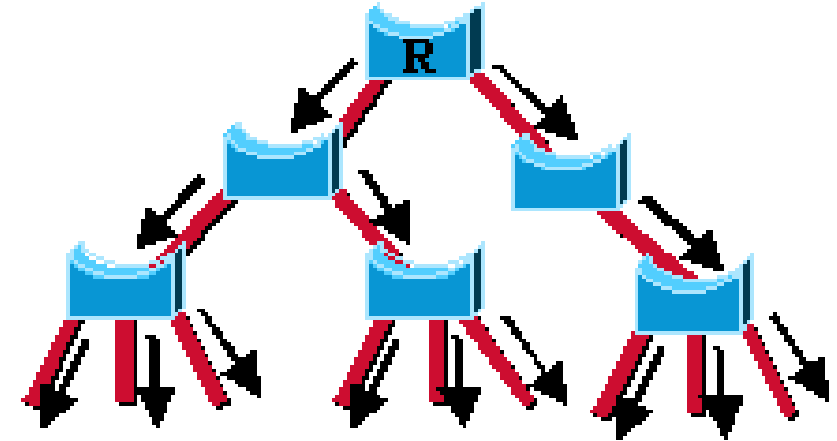
Proposal/Agreement Sequence



Mecanismos de mudança de topologia STP (802.1D)

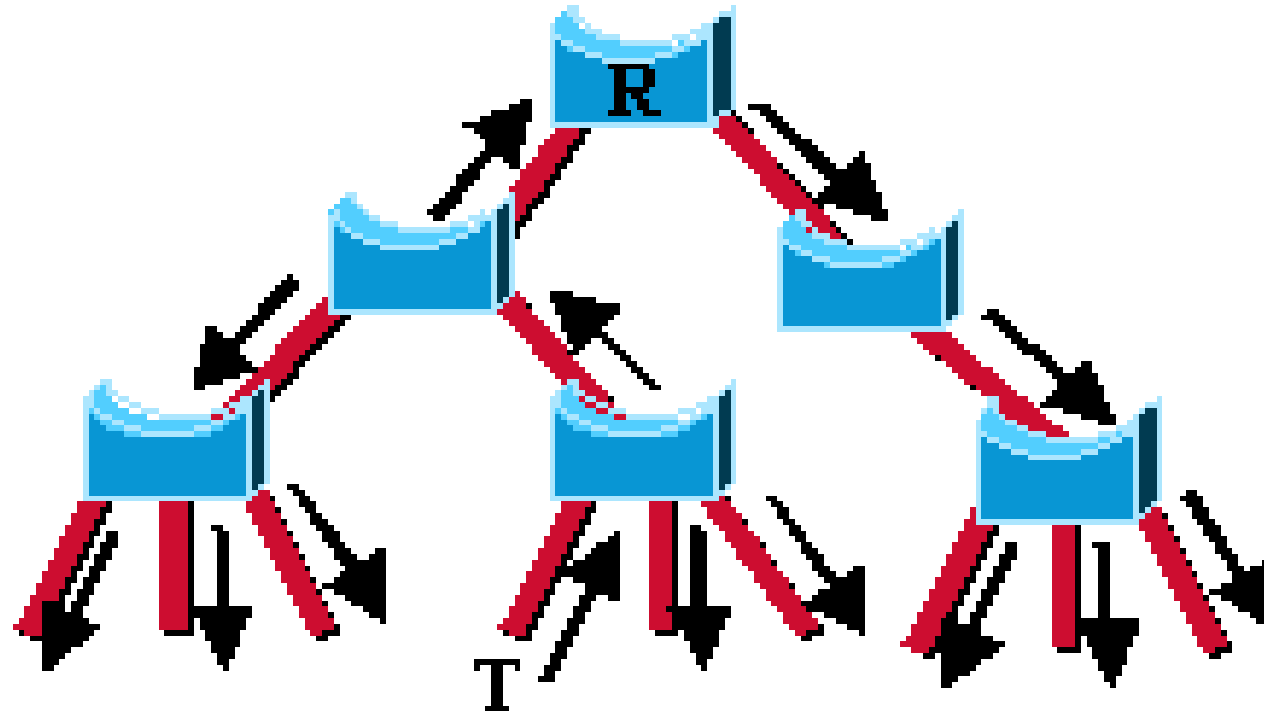


A topology change is generated on point T.
1st step: A TCN is going up to the root.



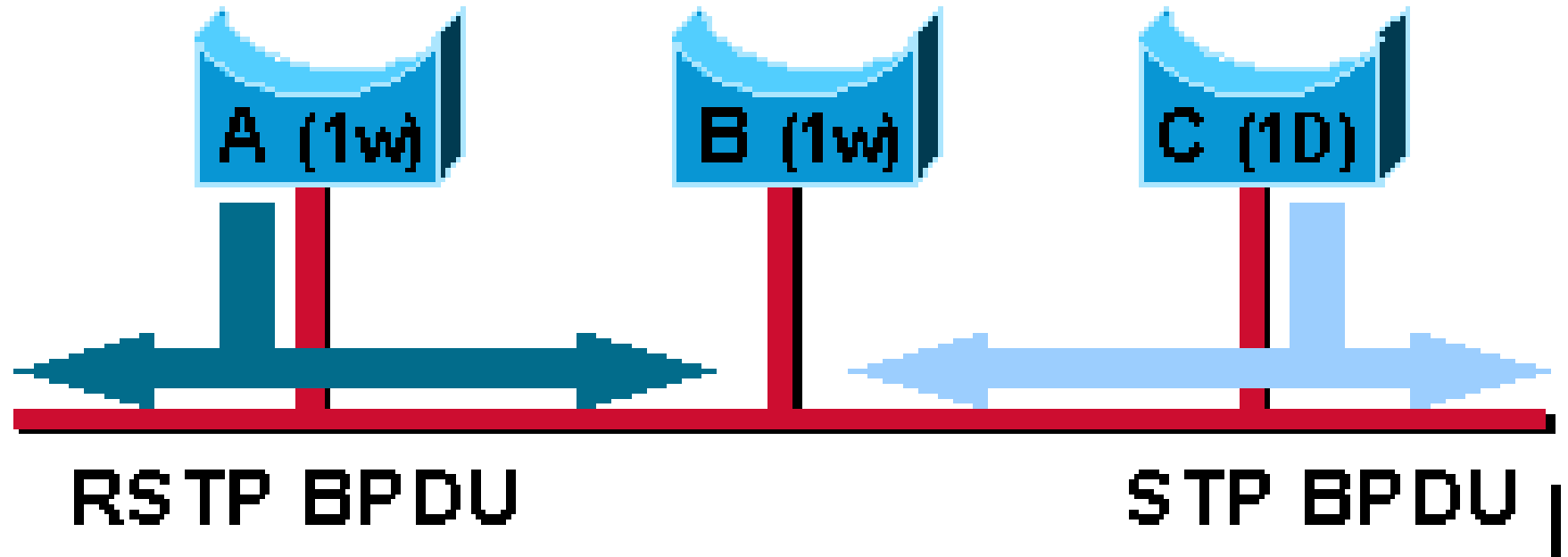
2nd step: the root advertises the TC for max-age+ forward delay.

Propagação da mudança de topologia

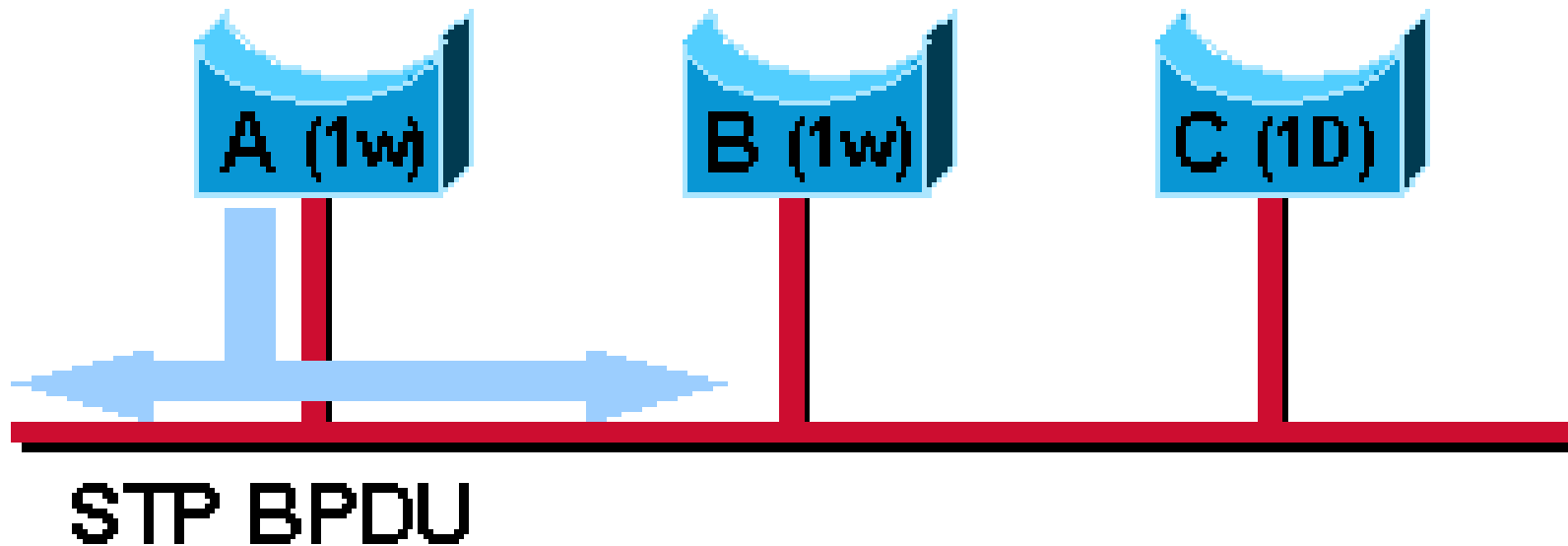


The originator of the TC directly floods this information through the network

Compatibilidade com 802.1D



Compatibilidade com 802.1D



PVST - *Per VLAN Spanning Tree*



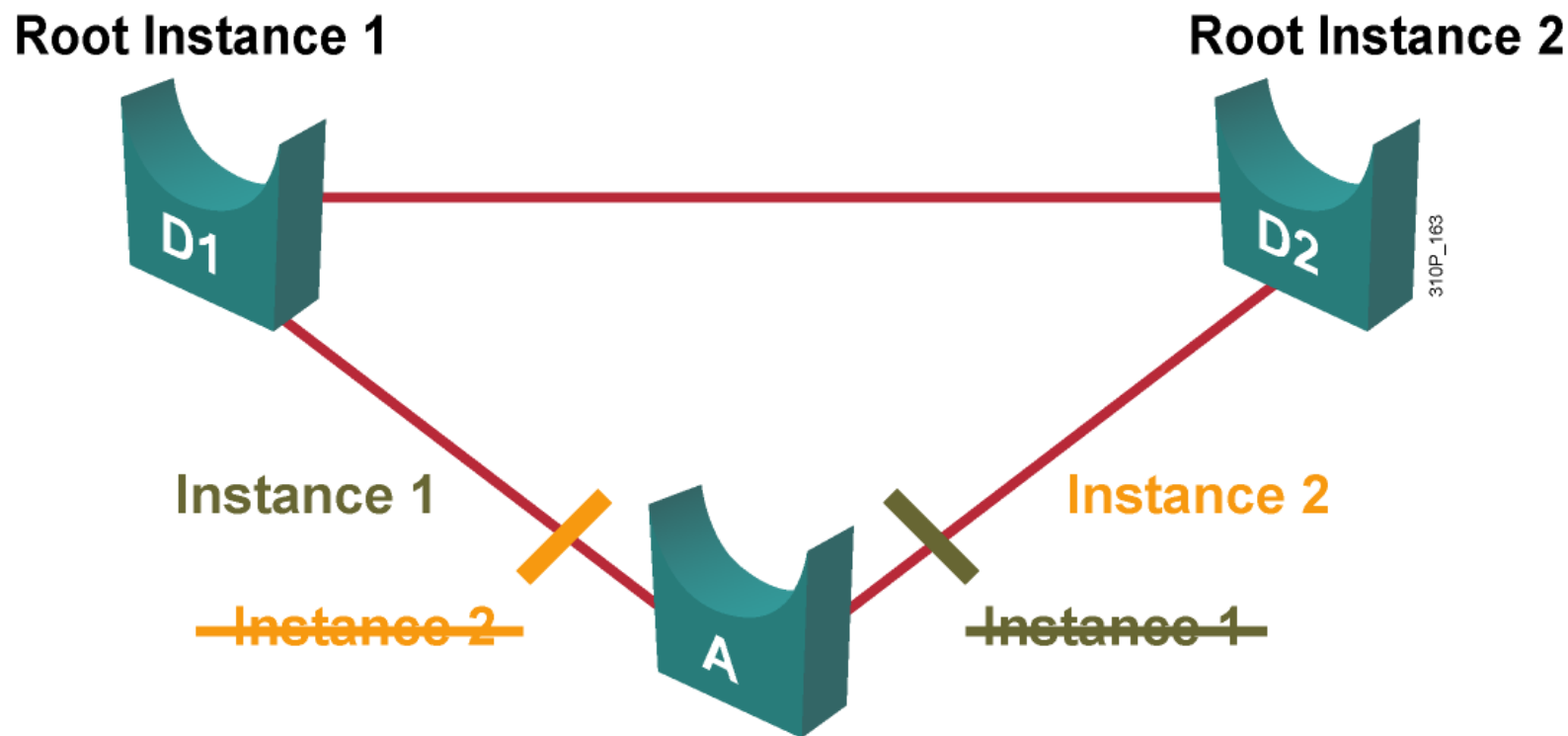
- Conceito
 - Como a topologia das VLAN pode ser diferente
 - Não basta ter um protocolo ST ao nível do porto físico
 - Uma porta pode estar *blocking* para evitar um *loop* numa VLAN e pode estar *forward* para outra VLAN
 - Um STP independente em cada VLAN
- Desvantagens
 - Não escala - num *switch* pode haver muitas VLAN (milhares)
 - Processamento é muito elevado para suportar os n STP independentes
 - Gera BPDU independentes para cada VLAN
 - Soluções: PVST+ e MSTP

MSTP - Multiple Spanning Tree Protocol



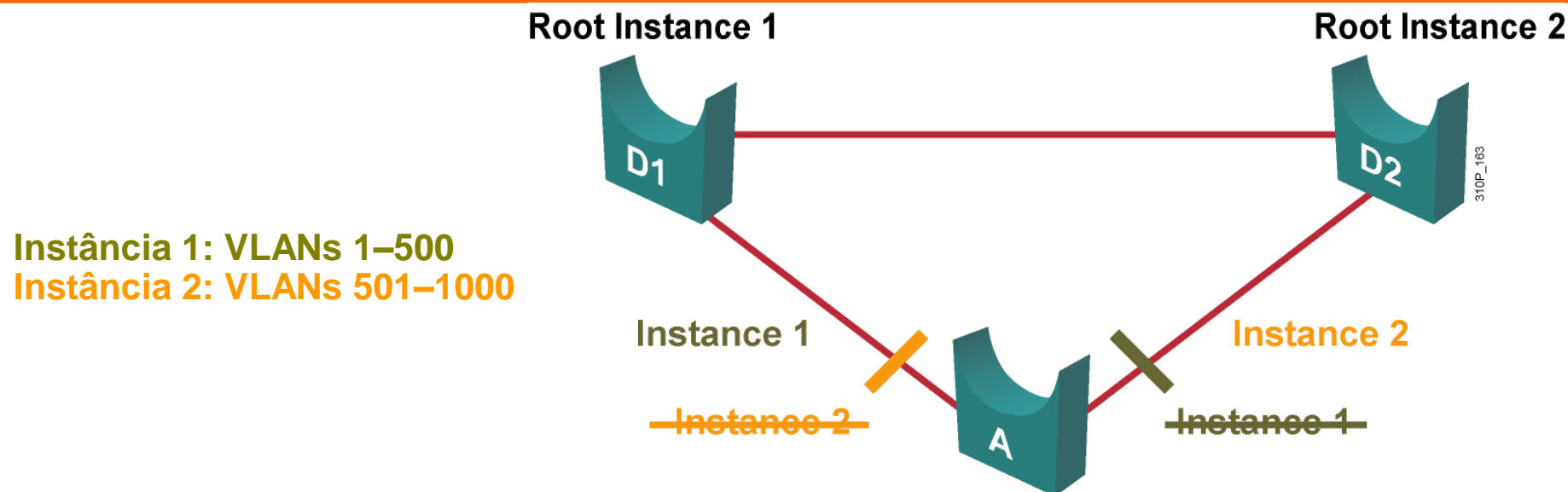
- Definido na norma IEEE 802.1s de Dezembro de 2002
- Objectivo:
 - Definir múltiplos STP cada um associado a um conjunto de VLAN
- Características:
 - A cada STP é associado um conjunto de VLAN com topologia semelhante

Multiple Spanning Tree Protocol – 802.1s



- Acesso do *switch* A, com 1000 VLAN ligadas em redundância, a dois *switches* de distribuição D1 e D2.
- Pretende-se que os utilizadores ligados ao *switch* A dividam o tráfego pelos dois *switches* D1 e D2.

Multiple Spanning Tree Protocol – 802.1s



- *Multiple Spanning Tree* (MST) estende o algoritmo RSTP IEEE 802.1w para múltiplas árvores de *spanning tree* (STP).
- O objetivo principal do MSTP é reduzir o número total de instâncias *spanning-tree* que está de acordo com a topologia física da rede e reduz o processamento dos *switches*.
- PVST+ considera uma instância STP para cada VLAN e não considera a topologia física da rede que pode não necessitar de tantas instâncias de STP diferentes.
- MSTP, por outro lado, usa o mínimo número de instâncias de STP que estão de acordo com a topologia física da rede.



- # Redes de Comunicação



- Sumário
 - *Rapid Spanning Tree Protocol (RSTP)*
 - *Per VLAN Spanning Tree Protocol (PVSTP)*
 - *Multiple Spanning Tree Protocol*